APR 10 1945

VOL. XXVIII JANUARY 1944 NO. 1

S
181
A45

THE JOURNAL OF AGRICULTURE

of the

UNIVERSITY OF PUERTO RICO

PAGE

A METHOD FOR THE SOLUTION OF NORMAL EQUATIONS
By Bernardo G. Capó 1

A METHOD OF INTERPRETING THE RESULTS OF FIELD TRIALS
By Bernardo G. Capó 7

A NEW METHOD OF PERFORMING FIELD TRIALS
By Bernardo G. Capó 22

PUBLISHED BY
THE AGRICULTURAL EXPERIMENT STATION
RÍO PIEDRAS, P. R.

AGRICULTURAL EXPERIMENT STATION

University of Puerto Rico — **Río Piedras, Puerto Rico**

JAIME BENÍTEZ, M.A., LL.M. ..... *Chancellor*
LUIS STÉFANI, M.S. ..... *Vice Chancellor*

STATION STAFF [1]

ARTURO ROQUE, M.S. ..... *Director*
R. COLÓN TORRES, M.S. ..... *Asst. Director*
E. DEL TORO, JR., M.S. [2] ..... *Admin. Asst.*
B. G. CAPÓ, PH.D. ..... *Biometrician*
M. R. GARCÍA ..... *Finance Officer*

Animal Husbandry:

————————, An. Husb., Head of Dept.
Carlos M. Muñiz, D.V.M., Assoc. Vet.
F. J. Marchán, M.S.,[2] Asst. Chem.
Herman Cestero, B.S., Actg. Asst. An. Husb.
Victor L. Quiñones, M.S., Actg. Asst. Chem.
Herbert R. Gómez, B.S.,[3] Actg. Asst. An. Husb.

Agronomy and Horticulture:

Arturo Riollano, M.S., Agron., Head of Dept.
Pedro Richardson, B.S., Agron.
Ferdinand Méndez, B.S., Assoc. Agron. in Charge of Station Farm
J. P. Rodríguez, B.S., Assoc. Agron.
Policarpo González, M.S., Actg. Assoc. Agron.
J. Guiscafré Arrillaga, M.S.,[2] Assoc. Agron.
E. Molinary Salés, B.S.,[3] Asst. Hort.
L. A. Gómez, B.S., Asst. Agron.
M. A. Díaz, B.S., Asst. Agron.
J. Lería Esmoris, B.S., Actg. Asst. Agron.

Chemistry:

Rafael Arroyo, B.S., S.E., Chem., Head of Dept.
H. Cruz Monclova, M.S., Ch.E.,[3] Assoc. Chem.
J. H. Ramírez, B.S., Assoc. Chem.
Leonardo Igaravídez, B.S., Asst. Chem.
Fernando Marrero, B.S., S.E., Asst. Chem.

Economics:

R. Colón Torres, M.S., Econ., Head of Dept.
L. M. Géigel, M.S., Assoc. Econ.
F. de Jesús, M.S., Assoc. Econ.
S. Díaz Pacheco, B.B.A.,[3] Assoc. Econ.
J. J. Serrallés, Jr., M.S., Assoc. Econ.
M. Roses, M.S.,[3] Asst. Econ.
M. Vélez, B.B.A., Asst. Econ.
M. Hernández, B.S., Actg. Asst. Econ.
J. O. Morales, B.S.,[3] Asst. Econ.
M. Piñero, B.S., Actg. Asst. Econ.
J. D. Rivera Anaya, B.S., Asst. Econ.
P. B. Vázquez Calcerrada, B.S., Actg. Asst. Econ.

Engineering:

M. L. Vicente, C.E., M.E., Engineer

Entomology:

G. N. Wolcott, Ph.D., Ent., Head of Dept.
L. F. Martorell, Ph.D., Assoc. Ent.
F. Soín, Jr., B.S., Assoc. Ent.

Phytopathology:

L. A. Alvarez, M.S., Phytopath., Head of Dept.
José Adsuar, B.S., Assoc. Phytopath.

Plant Breeding:

————————, Plant Breeder, Head of Dept.
A. Acosta Matienzo, M.S.,[2] Asst. Plant Breeder
Juan E. Veve, B.S., Actg. Asst. Plant Breeder
R. Vizcarrondo, B.S., Asst. Plant Breeder

Plant Physiology:

————————, Plant Physiol., Head of Dept.
F. J. Ramírez Silva, M.S., Assoc. Chem.
A. Gauthier Roig, B.S., Asst. Plant Physiol.

Soils:

J. A. Bonnet, Ph.D., Soil Chem., Head of Dept.
R. Rodríguez Torrent, B.S.,[3] Assoc. Soil Chem.
A. Riera, B.S., Asst. Chem.
P. Tirado Sulzona, B.S., Jr. Soil Conserv.
F. Mariota Trías, B.S., Jr. Soil Conserv.
G. M. Soto, B.S., Jr. Chem.

Substations:

Isabela:

————————, Agron. in Charge
C. J. Clavell, B.S.,[2] Asst. Agron.
R. K. Giles, B.S., Actg. Asst. Agron.
F. J. Juliá, B.S.,[2] Asst. Agron.
R. A. Irizarry, B.S., Actg. Asst. Agron.

Plant Propagation Substation:

J. S. Simons, B.S., Asst. Agron. in Charge

Animal Production Substation:

E. S. Vilella, B.S., Asst. An. Husb.

School of Tropical Medicine:

P. E. Serrano, B.S., Jr. Chem.
J. López Candal, B.S., Jr. Chem.

Library:

J. Rivera Ruíz, B.L.S., Librarian

---

[1] As of date of issue.
[2] On military leave.
[3] On leave.

Entered as second-class matter September 14, 1934, at the post office at Río Piedras, Puerto Rico, under the Act of August 24, 1912

181
AMS
38-30

Issued q
mation of

Vol. X

A M

The
been
the fu
viated
reduce
tions.
metho
classic
perien
"for t
matri
metho
attent
of co
metho

The
by Bô
"Pi
the *m*
each
The
tions
Let

A_{12}...

THE JOURNAL OF AGRICULTURE OF THE UNIVERSITY OF PUERTO RICO

Issued quarterly by the Agricultural Experiment Station of the University of Puerto Rico, for the publication of articles by members of its personnel, or others, dealing with any of the more technical aspects of scientific agriculture in Puerto Rico or the Caribbean Area.

Vol. XXVIII January, 1944 No. 1

A METHOD FOR THE SOLUTION OF NORMAL EQUATIONS

By BERNARDO G. CAPÓ

Biometrician

INTRODUCTION

The methods employed for the solution of a set of normal equations have been modified in recent years to employ modern calculating machines to the fullest extent. Remarkable among these modifications is the abbreviated Doolittle solution due to Dwyer (2), which requires a considerably reduced number of entries in the table of solution of such a system of equations. Dwyer (2, p. 443) claims that this method "is unquestionably the method which should be used by those seeking numerical solutions to the classical problems of multiple and partial correlation." Hotelling's experience with this and other methods has led him to state (4, p. 16) that "for the mass of least-square and other problems in which the inverse of a matrix is needed, the best procedure appears to begin with one of the methods described by Dwyer," etc. It is the purpose of this article to call attention to a method which is simpler, at least for relatively small numbers of constants to be fitted, than Dwyer's abbreviation of the Doolittle method.

THEORY OF THE METHOD

The method is based on Laplace's development of a determinant, stated by Bôcher (1, p. 26) as follows:

"Pick out any m rows (or columns) from a determinant D, and form all the m-rowed determinants from this matrix. The sum of the products of each of these minors by its algebraic complement is the value of D."

The application of this principle to the solution of a set of normal equations will be explained now.

Let

$$A_{123\ldots n} = \begin{vmatrix} a_1 & b_1 & c_1 & \cdots & m_1 \\ a_2 & b_2 & c_2 & \cdots & m_2 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ a_n & b_n & c_n & \cdots & m_n \end{vmatrix},$$

and let, for $i < j < k$, and with reference to $A_{123\ldots n}$,

$$A_{ij} = \begin{vmatrix} a_i & b_i \\ a_j & b_j \end{vmatrix} = a_i b_j - a_j b_i,$$

for example, $$A_{25} = \begin{vmatrix} a_2 & b_2 \\ a_5 & b_5 \end{vmatrix} = a_2b_5 - a_5b_2 ;$$

$$B_{ij} = \begin{vmatrix} b_i & c_i \\ b_j & c_j \end{vmatrix} = b_ic_j - b_jc_i ,$$

for example, $$B_{24} = \begin{vmatrix} b_2 & c_2 \\ b_4 & c_4 \end{vmatrix} = b_2c_4 - b_4c_2 ;$$

$$A_{ijk} = \begin{vmatrix} a_i & b_i & c_i \\ a_j & b_j & c_j \\ a_k & b_k & c_k \end{vmatrix} ; \quad \text{for example,} \quad A_{146} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_4 & b_4 & c_4 \\ a_6 & b_6 & c_6 \end{vmatrix} ;$$

$$A_{io j} = \begin{vmatrix} a_i & c_i \\ a_j & c_j \end{vmatrix} = a_ic_j - a_jc_i ,$$

for example, $$A_{206} = \begin{vmatrix} a_2 & c_2 \\ a_6 & c_6 \end{vmatrix} = a_2c_6 - a_6c_2 ; \text{ etc.}$$

Now, the application of the above-mentioned principle will show that

$$A_{123} = A_{12}c_3 - A_{13}c_2 + A_{23}c_1$$

$$A_{1234} = A_{123}d_4 - A_{124}d_3 + A_{134}d_2 - A_{234}d_1 , \text{ etc.}$$

The sign of any given term will be positive if "0" or an even number of inversions of adjacent subscripts is required to bring them into an ascending order, and negative if said number of inversions is odd.

Fisher's (3, p. 144) modification of Gauss' (5, p. 60 and 84) method of undetermined multipliers applied to the set of normal equations:

$$\begin{aligned} Aa_1 + Bb_1 + \cdots + Mm_1 &= Say, \\ Aa_2 + Bb_2 + \cdots + Mm_2 &= Sby, \\ \cdot \quad & \cdot \quad \cdot \\ \cdot \quad & \cdot \quad \cdot \\ \cdot \quad & \cdot \quad \cdot \\ \text{and} \qquad Aa_n + Bb_n + \cdots + Mm_n &= Smy, \end{aligned}$$

where $a_1 = Sa^2$, $b_1 = a_2 = Sab$, $\cdots$, $b_2 = Sb^2$, $c_2 = b_3 = Sbc$; etc., requires the solution of the following sets of equations for C_{aa}, C_{ab}, $\cdots$, C_{bb}, C_{bc}, $\cdots$, and C_{mm}:

$$\begin{aligned} \text{1st set:} \quad C_{aa}a_1 + C_{ab}b_1 + \cdots + C_{am}m_1 &= 1 \\ C_{aa}a_2 + C_{ab}b_2 + \cdots + C_{am}m_2 &= 0 \\ \cdot \quad & \cdot \quad \cdot \quad \cdot \\ \cdot \quad & \cdot \quad \cdot \quad \cdot \\ \cdot \quad & \cdot \quad \cdot \quad \cdot \\ C_{aa}a_n + C_{ab}b_n + \cdots + C_{am}m_n &= 0 \end{aligned}$$

$$\begin{aligned}
\text{2nd set: } & C_{ab}a_1 + C_{bb}b_1 + \cdots + C_{bm}m_1 = 0 \\
& C_{ab}a_2 + C_{bb}b_2 + \cdots + C_{bm}m_2 = 1 \\
& \cdot \quad \cdot \quad \cdot \quad \cdot \\
& \cdot \quad \cdot \quad \cdot \quad \cdot \\
& \cdot \quad \cdot \quad \cdot \quad \cdot \\
& C_{ab}a_n + C_{bb}b_n + \cdots + C_{bm}m_n = 0 \\
n\text{th set: } & C_{am}a_1 + C_{bm}b_1 + \cdots + C_{mm}m_1 = 0 \\
& C_{am}a_2 + C_{bm}b_2 + \cdots + C_{mm}m_2 = 0 \\
& \cdot \quad \cdot \quad \cdot \quad \cdot \\
& \cdot \quad \cdot \quad \cdot \quad \cdot \\
& \cdot \quad \cdot \quad \cdot \quad \cdot \\
& C_{am}a_n + C_{bm}b_n + \cdots + C_{mm}m_n = 1.
\end{aligned}$$

If one now lets $R = 1/A_{1234\ldots n}$, one can obtain the following relations for the c_{ij}'s.

From the 1st set of equations, one obtains

$$\begin{aligned}
C_{aa} &= B_{23\ldots n}R, \\
C_{ab} &= -A_{2034\ldots n}R, \\
C_{ac} &= A_{2504\ldots n}R, \\
&\cdot \\
&\cdot \\
&\cdot \\
C_{am} &= \pm A_{234\ldots n}R;
\end{aligned}$$

from the 2nd set, one obtains similarly,

$$\begin{aligned}
C_{bb} &= A_{1034\ldots n}R \\
C_{bc} &= -A_{1304\ldots n}R, \\
&\cdot \\
&\cdot \\
&\cdot \\
C_{bm} &= \pm_{134\ldots n}R;
\end{aligned}$$

and finally, from the nth set,

$$C_{mm} = A_{123\ldots(n-1)}R.$$

The values of the c_{ij}'s may then be found by obtaining the products of a number of determinants of the order $(n-1)$ by the reciprocal of $A_{123\ldots n}$, which itself may be evaluated from the relation

$$A_{123\ldots n} = A_{123\ldots(n-1)}M_n - A_{123\ldots(n-2)n}M_{n-1} + \cdots \pm A_{234\ldots n}M_1,$$

the last term being positive or negative according to the respective number of inversions of subscripts required as mentioned above. The determi-

nants which are coefficients of the m's in the last expression are the coefficients of R in the relations for C_{mm}, $C_{(m-1)m}$, $\cdots$, C_{am} above.

APPLICATION OF THE METHOD

The application of the method will be described in connection with the solution of a set of 3 normal equations involving 3 predictors.

Normal equations:

$$Aa_1 + Bb_1 + Cc_1 = Say,$$

$$Aa_2 + Bb_2 + Cc_2 = Sby,$$

$$Aa_3 + Bb_3 + Cc_3 = Scy.$$

From the previous relations,

$$C_{aa} = B_{23}R = (b_2c_3 - b_3c_2)R,$$

$$C_{ab} = -A_{203}R = -(a_2c_3 - a_3c_2)R,$$

$$C_{ac} = A_{23}R = (a_2b_3 - a_3b_2)R,$$

$$C_{bb} = A_{103}R = (a_1c_3 - a_3c_1)R,$$

$$C_{bc} = -A_{13}R = -(a_1b_3 - a_3b_1)R,$$

$$C_{cc} = A_{12}R = (a_1b_2 - a_2b_1)R,$$

$$A_{123} = A_{12}c_3 - A_{13}c_2 + A_{23}c_1,$$

and $R = 1/A_{123}$.

To evaluate the c_ij's, therefore, one needs to know A_{12}, A_{13}, A_{103}, A_{23}, A_{203}, and B_{23} to start with. The values of these expressions, which consist of differences between products of two terms each, as $pq - rs$, may be obtained from the machine without recording any intermediate figures. From these values and the original data, A_{123} and its reciprocal, R, may be evaluated, and from these, by setting R in the machine, the values of the c_ij's directly. The values of A, B and C and their standard errors may then be obtained in the usual way.

The application of the procedure outlined above to the same data from Carver used by Dwyer (2) to illustrate his method follows.

$a_1 = 1.000$	$b_1 = 0.313$	$c_1 = 0.280$	$Say = 0.495$
$a_2 = 0.313$	$b_2 = 1.000$	$c_2 = 0.652$	$Sby = 0.650$
$a_3 = 0.280$	$b_3 = 0.652$	$c_3 = 1.000$	$Scy = 0.803$
			$Syy = 1.000$

70 696 AA A 30

$$A_{12} = a_1b_2 - a_2b_1 = 0.902031 \quad C_{cc} = A_{12}R = 1.758997$$
$$A_{13} = a_1b_3 - a_3b_1 = 0.564360 \quad C_{bc} = -A_{13}R = -1.100525$$
$$A_{103} = a_1c_3 - a_3c_1 = 0.921600 \quad C_{bb} = A_{103}R = 1.797157$$
$$A_{23} = a_2b_3 - a_3b_2 = -0.075924 \quad C_{ac} = A_{23}R = -0.148055$$
$$A_{203} = a_2c_3 - a_3c_2 = 0.130440 \quad C_{ab} = -A_{203}R = -0.254363$$
$$B_{23} = b_2c_3 - b_3c_2 = 0.574896 \quad C_{aa} = B_{23}R = 1.121070$$
$$A_{123} = A_{12}c_3 - A_{13}c_2 + A_{23}c_1 = 0.512810$$
$$R = 1/A_{123} = 1.950040$$
$$A = C_{aa}Say + C_{ab}Sby + C_{ac}Scy = 0.270706$$
$$B = C_{ab}Say + C_{bb}Sby + C_{bc}Scy = 0.158521$$
$$C = C_{ac}Say + C_{bc}Sby + C_{cc}Scy = 0.623846$$

Reduced $Syy = 1 - ASay - BSby - CScy = 0.262014$.

The information contained in rows 4 to 10 inclusive, of table VII, of Dwyer's article (2), which required *37* entries (disregarding the zeros), is found in the last *12* rows of the preceding scheme, which contains but *18* entries. In addition, the preceeding scheme contains the values of C_{ac}, C_{ab}, and C_{bc}, which do not appear in said table VII, but are essential for the statistical evaluation of differences between A, B, and C. The calculation of these values would require still 3 more entries in that table, raising the total number of entries to 40.

For the case of 4 normal equations with 4 predictors the respective numbers of entries would be *63* in the abbreviated Doolittle solution method and *40* in the proposed method. Two additional advantages of the proposed method as against Dwyer's abbreviated Doolittle solution are that the operations tend to be simpler and errors tend to be of less importance in the proposed method, since an error will not invalidate such a large proportion of the work performed subsequently to it as it does in the Doolittle technique when no check columns are included.

SUMMARY

A new method of solving normal equations is presented. A comparison is made of the application of the method to a multiple regression case used by Dwyer to illustrate his abbreviation of the Doolittle method. The proposed method for relatively small numbers of unknowns at least, is simpler than Dwyer's method, considered to date to be one, if not the most efficient for the purpose.

REFERENCES

1. Bôcher, M., *Introduction to Higher Algebra*, The Macmillan Co., New York, 1907.
2. Dwyer, P. S., Recent developments in correlation technique, *J. Am. Stat. Asso.*, **37** (220), 441–460. 1942.

3. Fisher, R. A., *Statistical Methods for Research Workers*, 4th Ed., Oliver and Boyd, Edinburgh, 1932.
4. Hotelling, H., Some new methods in matrix calculation, *Ann. Math. Stat.*, 14 (1), 1–34. 1943.
5. Merriman, M. *Method of Least Squares*, 8th ed., revised, John Wiley and Sons, Inc., New York, 1911.

A METHOD OF INTERPRETING THE RESULTS OF FIELD TRIALS

By BERNARDO G. CAPÓ

Biometrician

In the comparison by means of field trials of a relatively small number of agricultural practices, the complete randomized block and Latin square designs are most commonly employed by the majority of research workers. When a relatively large number of such practices are to be tested, other designs are used, such as the incomplete randomized blocks, including the lattices, and designs which make use of check plots in one form or another. The reason for this attitude is that the effects of the differences in fertility between the different plots within any given block, file or column, as the case may be, are not eliminated from the estimate of the experimental error obtained by interpreting tests performed according to the above-mentioned designs by the simple methods of statistical analysis developed for use with said types of experiments.

It occasionally happens, however, that field tests are performed in places where the fertility of the soil varies a great deal between spots relatively near to one another. Under these conditions, and especially with crops in connection with which relatively large plots must be used, the usual method of statistical analysis yields too high an estimate of error and it is impossible to determine any significant difference between the practices under trial. The possibility of the use of some alternate method of interpreting the results of such experiments might conceivably be of benefit in a number of cases. It is the purpose of this article to suggest the use of a method applicable to cases of this kind which occur when relatively small numbers of practices have been tested.

The method is based on the assumption of a different effect constant for every different pair of adjacent plots in the field when an even number of plots has been used. If an odd number of plots has been used, the previous assumption is made for all but 3 plots lying together in the field, for all three of which the same effect constant is assumed to hold. By fitting then a multiple regression equation to the results of such an experiment, it is possible to obtain an estimate of the effect constant of each of the practices tested. Mathematical statisticians certainly need no more information with respect to the way in which the method works than the one conveyed in the three preceding sentences. The rank and file of field workers, however, are not so fortunate in this respect. Due to this, the explanation of the method will be performed by following a numerical example throughout for a case where an even number of plots was used. The slight modi-

fication needed for the cases of odd numbers of plots will be dealt with further on.

DIAGRAM I

Field arrangement of alfalfa experiment

50	40	30	20	10
A	H	G	F	E
30	26	30	35	44
49	39	29	19	9
C	J	I	H	G
37	36	29	30	34
48	38	28	18	8
E	B	A	J	I
59	38	22	29	33
47	37	27	17	7
I	F	E	D	C
60	42	35	24	30
46	36	26	16	6
G	D	C	B	A
62	52	45	38	33
45	35	25	15	5
D	A	J	I	H
57	33	39	40	48
44	34	24	14	4
H	E	D	C	B
43	38	43	39	49
43	33	23	13	3
E	I	H	G	F
39	36	41	45	57
42	32	22	12	2
F	C	B	A	J
43	42	47	40	52
41	31	21	11	1
J	G	F	E	D
48	46	50	52	56

Diagram I shows the field arrangement of the plots of an alfalfa experiment performed at the Agricultural Experiment Substation farm at Isabela by Messrs. L. A. Serrano and C. J. Clavell. In each case, the first number

is the plot identification number, the letter beneath it is the treatment identification letter, and the number beneath it is the observed yield in tons of green alfalfa roughage per acre obtained from 16 successive cuttings.

In table I the treatments are described, and in addition, the total and mean yields obtained with each treatment are presented.

Although in this experiment ten different treatments were tested, each one of the treatments may be expressed in terms of the amounts of nitrogen, phosphorus and potassium applied per acre. Thus, instead of trying to determine whether the observed difference in mean yields caused under the effects of any two given treatments is significant or not, one may try to determine whether the nitrogen applications have affected the alfalfa yields

TABLE I

Yields in tons of green alfalfa roughage per acre

Letter	Treatments						Yields	
	Units applied			Lbs./A.				
	N	P	K	NH_3	P_2O_5	K_2O	Total	Mean
A	0	0	0	0	0	0	158	31.6
B	0	1	3	0	48	216	211	42.2
C	0	1	4	0	48	288	193	38.6
D	1	1	3	36	48	216	232	46.4
E	2	1	3	72	48	216	228	45.6
F	1	1	2	36	48	144	227	45.4
G	1	1	4	36	48	288	217	43.4
H	1	0	3	36	0	216	188	37.6
I	1	2	3	36	96	216	198	39.6
J	2	2	4	72	96	288	204	40.8

significantly or not; and similarly for the other two elements. In order to determine the effects of these elements, one may assume that the yield of any plot is the additive result of a series of terms as follows:

$$N_i b_n + P_i b_p + K_i b_k + B_i = Y_i , \qquad (1)$$

where Y_i is the yield of plot i; B_i is the yield constant of the block of which plot i forms part; N_i, P_i, and K_i are the units of nitrogen, phosphorus and potassium applied to the crop in plot i; and b_n, b_p, and b_k are the respective increases in yield caused by the application of each unit of nitrogen, phosphorus or potassium respectively. b_n, b_p, and b_k are termed partial regression coefficients. The values of these coefficients are then the ones to be tested to determine whether the applications of each respective element have affected the yields significantly or not.

Thus, if one assumes block 1 to consist of plots 1 and 2, and assumes

further that the yield constant of the plots in block 1 is B_1, one may write the following equation for plot 1:

$$b_n + b_p + 3b_k + B_1 = 56, \tag{2}$$

since plot 1 received treatment D, which consisted in the application of 1 unit of nitrogen, 1 unit of phosphorus and 3 units of potassium. One may similarly write the following equation for plot 2:

$$2b_n + 2b_p + 4b_k + B_1 = 52. \tag{3}$$

In a similar way, one may write down equations for the other plots, obtaining a total of 50 equations, one for each plot. In these equations there would be the 28 unknowns b_n, b_p, b_k, B_1, B_2, $\cdots$, B_{25}. It is impossible to determine values for these 28 unknowns which will fit the 50 equations. The next best solution, that of finding the most probable values of the constants, must be resorted to. Once these most probable values are determined, one may estimate the yield of any plot by means of equation (1) above. The difference between the actually observed yield on any given plot and the value of said yield estimated by means of equation (1) is the error of estimate of the yield of that plot. Thus, for plot (1) the error of estimate would be:

$$d_1 = 56 - b_n - b_p - 3b_k - B_1, \tag{4}$$

since the actually observed yield was 56 and the estimated yield would be the value of $b_n + b_p + 3b_k + B_1$ when the most probable values of these constants are substituted in equation (4).

According to the principle of Least Squares, the most probable values of constants like the ones under discussion are those which render the sum of the squares of the errors of estimate a minimum. Thus, if d_1 is the error of estimate of plot 1, d_2 is the error of estimate of plot 2, etc., those values of the constants b_n, b_p, b_k, B_1, B_2, $\cdots$, B_{25} which would make $Sd^2 = d_1^2 + d_2^2 + \cdots + d_{50}^2$ a minimum, would be their most probable values, and therefore, the ones to be used in equation (1) for estimating the yield of any given plot.

Now,

$$d_1^2 = (56 - b_n - b_p - 3b_k - B_1)^2. \tag{5}$$

Similarly,

$$d_2^2 = (52 - 2b_n - 2b_p - 4b_k - B_1)^2. \tag{6}$$

In a similar way one may find the values of d_3^2, d_4^2, $\cdots$, d_{50}^2.

In applying the criteria for a minimum to the expression $Sd^2 = d_1^2 + d_2^2 + \cdots + d_{50}^2 = (56 - b_n - b_p - 3b_k - B_1)^2 + (52 - 2b_n - 2b_p -$

$4b_k - B_1)^2 + \cdots + (30 - B_{25})^2$, one obtains 28 equations, each one of which is a partial derivative of Sd^2 with respect to one of the constants b_n, b_p, b_k, B_1, B_2, $\cdots$, B_{25} equated to zero. The equation obtained on finding the partial derivative of Sd^2 with respect to B_1 and simplifying is

$$56 - b_n - b_p - 3b_k - B_1 + 52 - 2b_n - 2b_p - 4b_k - B_1 = 0. \quad (7)$$

From this equation one gets

$$B_1 = 54 - 3b_n/2 - 3b_p/2 - 7b_k/2. \quad (8)$$

If one now substitutes this value of B_1 in equation (4) above, and simplifies, one gets

$$d_1 = 2 + b_n/2 + b_p/2 + b_k/2 = (d_1 - d_2)/2. \quad (9)$$

Substituting the value of B_1 in the corresponding equation for plot 2, i.e.,

$$d_2 = 52 - 2b_n - 2b_p - 4b_k - B_1,$$

one gets $$d_2 = -2 - b_n/2 - b_p/2 - b_k/2 = -(d_1 - d_2)/2. \quad (10)$$

It will be noticed that by means of these substitutions B_1 has been eliminated from the equations corresponding to the first two plots and, since these are the only two equations where B_1 occurs, from the whole set of 50 equations. In a similar way all the other B's may be eliminated from the other equations resulting finally in a system of 50 equations in but the three unknowns b_n, b_p, and b_k. In this new system of equations, the right-hand side of the equation corresponding to plot 1 will be equal to the right-hand side of that corresponding to plot 2 except that the signs are reversed. Similarly the right-hand sides of the equations corresponding to plots 3 and 4 will be equal except for the signs changed; and similarly for the remaining pairs of equations.

Now, the most probable values of the three unknowns b_n, b_p, and b_k are, as stated above, those which would make Sd^2 a minimum. From equations (9) and (10) above, it follows that $d_1 = -d_2$, and therefore, $d_1^2 = d_2^2$ and $-2d_1d_2 = 2d_1^2 = 2d_2^2 = d_1^2 + d_2^2$. Thus, $(d_1 - d_2)^2 = d_1^2 - 2d_1d_2 + d_2^2 = 2d_1^2 + 2d_2^2$. A similar relation may be found in the cases of the corresponding errors of estimate of the plots belonging to the other blocks. Therefore,

$$(d_1 - d_2)^2 + (d_3 - d_4)^2 + \cdots + (d_{49} - d_{50})^2 = 2d_1^2 + 2d_2^2 + 2d_3^2 + 2d_4^2 + \cdots + 2d_{49}^2 + 2d_{50}^2 = 2Sd^2. \quad (11)$$

For simplicity of calculations Sd^2 is calculated in the method under discussion by calculating $2Sd^2$ by the use of the left-hand side of equation (11), and dividing by 2. The numerical values of the coefficients of the

squares and products of the unknowns which appear in Sd^2 may thus be conveniently found in the example under study by arranging the calculations as they appear in table II.

TABLE II

Calculation of the numerical values of the coefficients of squares and products

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Difference															
Plots	Treatments	d_N	d_P	d_K	d_Y	d_N^2	$d_N d_P$	$d_N d_K$	$d_N d_Y$	d_P^2	$d_P d_K$	$d_P d_Y$	d_K^2	$d_K d_Y$	d_Y^2
1–2	D-J	−1	−1	−1	4	1	1	1	−4	1	1	−4	1	−4	16
3–4	F-B	1	0	−1	8	1	0	−1	8	0	0	0	1	−8	64
5–6	H-A	1	0	3	15	1	0	3	15	0	0	0	9	45	225
8–7	I-C	1	1	−1	3	1	1	−1	3	1	−1	3	1	−3	9
10–9	E-G	1	0	−1	10	1	0	−1	10	0	0	0	1	−10	100
11–12	E-A	2	1	3	12	4	2	6	24	1	3	12	9	36	144
13–14	G-C	1	0	0	6	1	0	0	6	0	0	0	0	0	36
15–16	I-B	1	1	0	2	1	1	0	2	1	0	2	0	0	4
18–17	J-D	1	1	1	5	1	1	1	5	1	1	5	1	5	25
20–19	F-H	0	1	−1	5	0	0	0	0	1	−1	5	1	−5	25
21–22	F-B	1	0	−1	3	1	0	−1	3	0	0	0	1	−3	9
24–23	D-H	0	1	0	2	0	0	0	0	1	0	2	0	0	4
26–25	C-J	−2	−1	0	6	4	2	0	−12	1	0	−6	0	0	36
27–28	E-A	2	1	3	13	4	2	6	26	1	3	13	9	39	169
30–29	G-I	0	−1	1	1	0	0	0	0	1	−1	−1	1	1	1
31–32	G-C	1	0	0	4	1	0	0	4	0	0	0	0	0	16
34–33	E-I	1	−1	0	2	1	−1	0	2	1	0	−2	0	0	4
36–35	D-A	1	1	3	19	1	1	3	19	1	3	19	9	57	361
37–38	F-B	1	0	−1	4	1	0	−1	4	0	0	0	1	−4	16
39–40	J-H	1	2	1	10	1	2	1	10	4	2	20	1	10	100
41–42	J-F	1	1	2	5	1	1	2	5	1	2	5	4	10	25
44–43	H-B	1	−1	0	4	1	−1	0	4	1	0	−4	0	0	16
46–45	G-D	0	0	1	5	0	0	0	0	0	0	0	1	5	25
47–48	I-E	−1	1	0	1	1	−1	0	−1	1	0	1	0	0	1
49–50	C-A	0	1	4	7	0	0	0	0	1	4	7	16	28	49
Sum.....						29	11	18	133	20	16	77	67	199	1480
Coefficients (= Sum/2).....						14.5	5.5	9	66.5	10	8	38.5	33.5	99.5	740

Columns (1), (2) and (6) have been filled from the information given in diagram I. In finding the differences corresponding to any given pair of adjacent plots, said differences have been taken in such an order as to yield a positive value of d_Y in every case. Thus, since the yield of plot 2 was smaller than the yield of plot 1, the differences are to be found by subtracting the data of plot 2 from the corresponding data of plot 1. This is indicated by writing "1 − 2" in column (1). Since plot 1 received treatment *D*

and plot 2 received treatment J, the corresponding entry in column (2) is $D - J$. The entry in column (6) is $4(= 56 - 52)$. The rest of columns (1), (2), and (6) is filled in a similar way. Thus, for plots 33 and 34, the order will be $34 - 33$ since the yield of plot 34 exceeded the yield of plot 33. In columns (2) and (6) the corresponding entries are $E - I$ and 2, respectively.

After filling columns (1), (2), and (6); columns (3), (4), and (5) must be filled. In filling them use is made of table I. The entries corresponding to plots 1 and 2 in these columns are found as follows: d_N = units of N of treatment D minus the units of N of treatment $J = 1 - 2 = -1$; and likewise, $d_P = 1 - 2 = -1$, and $d_K = 3 - 4 = -1$. For plots 33 and 34, where the corresponding difference between treatments is $E - I$, $d_N = 2 - 1 = 1$, $d_P = 1 - 2 = -1$, and $d_K = 3 - 3 = 0$. In a similar way, the rest of columns (3), (4), and (5) may be filled.

Once columns (1) to (6) are filled, the corresponding entries in columns (7) to (16) may be made by performing in each case the operation indicated by the heading at the top of each column. Thus, for the entries corresponding to plots 1 and 2, $d_N^2 = d_N d_N = (-1)(-1) = 1$; $d_N d_P = (-1)(-1) = 1$; $d_N d_K = (-1)(-1) = 1$; $d_N d_Y = (-1)(4) = -4$; $d_P^2 = d_P d_P = (-1)(-1) = 1$; $d_P d_K = (-1)(-1) = 1$; $d_P d_Y = (-1)\,4 = -4$; etc. The entries corresponding to plots 33 and 34 are, likewise, as follows: $d_N^2 = (1)(1) = 1$; $d_N d_P = (1)(-1) = -1$; $d_N d_K = (1)(0) = 0$; $d_N d_Y = (1)(2) = 2$; etc.

Once these entries are made, it remains but to add the entries of columns (6) to (16) and to divide each sum by 2. Since

$$Sd^2 = b_n^2 Sn^2 + 2b_n b_p Snp + 2b_n b_k Snk - 2b_n Sny + b_p^2 Sp^2 + 2b_p b_k Spk - 2b_p Spy + b_k^2 Sk^2 - 2b_k Sky + Sy^2, \quad (12)$$

one can write the value of Sd^2 for this case as follows:

$$Sd^2 = 14.5b_n^2 + 2(5.5)b_n b_p + 2(9)b_n b_k - 2(66.5)b_n + 10b_p^2 + 2(8)b_p b_k - 2(38.5)b_p + 33.5b_k^2 - 2(99.5)b_k + 740. \quad (13)$$

As previously stated, the most probable values of b_n, b_p, and b_k are those which will make Sd^2 minimum. Among the mathematical requirements for Sd^2 to be a minimum, the requirements stated by three following equations must be fulfilled:

$$b_n Sn^2 + b_p Snp + b_k Snk = Sny, \quad (14)$$

$$b_n Snp + b_p Sp^2 + b_k Spk = Spy, \quad (15)$$

and

$$b_n Snk + b_p Spk + b_k Sk^2 = Sky. \quad (16)$$

Equations (14), (15), and (16) are known as "normal equation in b_n," "normal equation in b_p," and "normal equation in b_k," respectively. In

the sample under study, the equations are as follows:

$$14.5b_n + 5.5b_p + 9b_k = 66.5, \tag{17}$$

$$5.5b_n + 10b_p + 8b_k = 38.5, \tag{18}$$

and

$$9b_n + 8b_p + 33.5b_k = 99.5. \tag{19}$$

These three simultaneous linear equations in 3 unknowns may be solved by any of the well-known methods studied in the elementary courses in algebra to obtain the values of the partial regression coefficients b_n, b_p, and b_k. Due, however, to the fact that the knowledge of the standard error of any given statistic is almost as essential as the value of the statistic itself for the proper evaluation of the significance of said statistic and of differences between it and some other statistic, Fisher's modification of Gauss' method of correlatives or indeterminate multipliers is usually resorted to in practice. Snedecor (4, p. 302) and Rider (3, p. 39) present discussions and give examples of solutions of these systems of equations by this method. Doolittle's method of solution of normal equations as discussed by Mills (2, p. 656) may be incorporated to advantage to the method of indeterminate multipliers. The author has found these methods, however, too tedious and intricate for use in practice. He suggests, therefore, the use of the following method which, for a small number of constants to be fitted, is much easier and takes considerably less time to apply than the methods previously referred to, specially if one has a printed or mimeographed form stating the calculations to be performed, so that one has to fill in merely the values asked for in said form. The suggested method of solution is as follows:

Following Fisher, (1, p. 144), the equations to be solved are:

$$C_{aa}a_1 + C_{ab}b_1 + C_{ac}C_1 = 1, \tag{20}$$

$$C_{aa}a_2 + C_{ab}b_2 + C_{ac}C_2 = 0, \tag{21}$$

$$C_{aa}a_3 + C_{ab}b_3 + C_{ac}C_3 = 0, \tag{22}$$

$$C_{ab}a_1 + C_{bb}b_1 + C_{bc}C_1 = 0, \tag{23}$$

$$C_{ab}a_2 + C_{bb}b_2 + C_{bc}C_2 = 1, \tag{24}$$

$$C_{ab}a_3 + C_{bb}b_3 + C_{bc}C_3 = 0, \tag{25}$$

$$C_{ac}a_1 + C_{bc}b_1 + C_{cc}C_1 = 0, \tag{26}$$

$$C_{ac}a_2 + C_{bc}b_2 + C_{cc}C_2 = 0, \tag{27}$$

and

$$C_{ac}a_3 + C_{bc}b_3 + C_{cc}C_3 = 1, \tag{28}$$

where $a_1 = Sn^2$, $a_2 = b_1 = Snp$, $a_3 = c_1 = Snk$, $b_2 = Sp^2$, $b_3 = c_2 = Spk$, and $c_3 = Sk^2$.

Now, if A_{12} represents the determinant of the second order $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$, A_{123} means the determinant of the third order $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$, and, in general, $A_{123\cdots n}$ means the determinant of the nth order

$$\begin{vmatrix} a_1 & b_1 & \cdots & m_1 \\ a_2 & b_2 & \cdots & m_2 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_n & b_n & \cdots & m_n \end{vmatrix},$$

the following relations may be easily deducted from the fact that any given determinant is equal to the sum of all terms which may be formed by writing down the principal diagonal $a_1 b_2 c_3 \cdots m_n$ and forming all possible permutations of the subscripts; the sign of each term being positive if an even number of inversions must be performed to bring the subscripts into an ascending series and negative if said number of inversions is odd.

$$A_{12} = a_1 b_2 - a_2 b_1 ; \qquad A_{123} = A_{12} c_3 - A_{13} c_2 + A_{23} c_1 ,$$

$$A_{1234} = A_{123} d_4 - A_{124} d_3 + A_{134} d_2 - A_{234} d_1 , \text{ etc.}$$

If now $A_{103} = a_1 c_3 - a_3 c_1$; $A_{203} = a_2 c_3 - a_3 c_2$ and $B_{23} = b_2 c_3 - b_3 c_2$, the values of C_{aa} , C_{ab} , and C_{ac} may be determined from equations (20), (21), and (22) above to be as follows:

$$C_{aa} = B_{23} A_{123}^{-1} ; \quad C_{ab} = -A_{203} A_{122}^{-1} ; \quad \text{and} \quad C_{ac} = A_{23} A_{123}^{-1} .$$

From equations (23), (24), and (25), the values of C_{bb} and C_{bc} come out to be as follows:

$$C_{bb} = A_{103} A_{123}^{-1} ; \quad \text{and} \quad C_{bc} = -A_{13} A_{123}^{-1} .$$

Finally, from equations (26), (27), and (28),

$$C_{cc} = A_{12} A_{123}^{-1} .$$

The number of operations required to evaluate these statistics is not large and does not take too long, specially if, as already stated, one has some form indicating the operations and has to fill only the numbers in it. Such a form, already filled, appears as Table III.

In this table, A corresponds to b_n , B corresponds to b_p , and C corresponds

to b_k which are the unknowns whose values were to be determined. These values of b_n , b_p , and b_k may be checked by substituting them in equations (17), (18), and (19) above.

TABLE III

Solution of simultaneous linear equations

$$Aa_1 + Bb_1 + Cc_1 = Say$$
$$Aa_2 + Bb_2 + Cc_2 = Sby$$
$$Aa_3 + Bb_3 + Cc_3 = Scy$$

$a_1 = 14.5$	$b_1 = 5.5$	$c_1 = 9$	$Say = 66.5$
$a_2 = 5.5$	$b_2 = 10$	$c_2 = 8$	$Sby = 38.5$
$a_3 = 9$	$b_3 = 8$	$c_3 = 33.5$	$Scy = 99.5$
$a_1b_2 = 145$	$a_2b_1 = 30.25$	$a_3b_1 = 49.5$	$b_2c_3 = 335$
$a_1b_3 = 116$	$a_2b_3 = 44$	$a_3b_2 = 90$	$b_3c_2 = 64$
$a_1c_3 = 485.75$	$a_2c_3 = 184.25$	$a_3c_1 = 81$	
		$a_3c_2 = 72$	

$A_{12} = a_1b_2 - a_2b_1 = 114.75$ $\quad A_{12}c_3 = 3844.125$

$A_{13} = a_1b_3 - a_3b_1 = 66.50$ $\quad -A_{13}c_2 = -532$

$A_{123} = a_1c_3 - a_3c_1 = 404.75$ $\quad A_{23}c_1 = -414$

$A_{23} = a_2b_3 - a_3b_2 = -46$ $\quad A_{123} = 2898.125$

$A_{203} = a_2c_3 - a_3c_2 = 112.25$ $\quad A^{-1}_{123} = 0.000345051$

$B_{23} = b_2c_3 - b_3c_2 = 271$

$C_{aa} = B_{23}A^{-1}_{123} = 0.0935088$ $\quad C_{bb} = A_{103}A^{-1}_{123} = 0.139659$

$C_{ab} = -A_{203}A^{-1}_{123} = -0.0387320$ $\quad C_{bc} = -A_{13}A^{-1}_{123} = -0.0229459$

$C_{ac} = A_{23}A^{-1}_{123} = -0.0158723$ $\quad C_{cc} = A_{12}A^{-1}_{123} = 0.0395946$

Checks:

$C_{aa}a_1 = 1.35588$	$C_{ab}a_2 = -0.213026$	$C_{ac}a_3 = -0.142851$
$C_{ab}b_1 = -0.213026$	$C_{bb}b_2 = 1.39659$	$C_{bc}b_3 = -0.183567$
$C_{ac}c_1 = -0.142851$	$C_{bc}c_2 = -0.183567$	$C_{cc}c_3 = 1.32642$
1.0000	1.0000	1.0000

$C_{aa}Say = 6.218335$	$C_{ab}Say = -2.575678$	$C_{ac}Say = -1.055508$
$C_{ab}Sby = -1.491182$	$C_{bb}Sby = 5.376872$	$C_{bc}Sby = -0.883417$
$C_{ac}Scy = -1.579294$	$C_{bc}Scy = -2.283117$	$C_{cc}Scy = 3.939663$
$A = 3.147859$	$B = 0.518077$	$C = 2.000738$

The significance of these coefficients of regression may be checked by calculating their standard errors and using the "*t*-test." The total sum of squared deviations corrected for variations in fertility from block to block was found in table II to be 740, subject to 25 degrees of freedom, since from the total original number of 49 degrees of freedom, 24 (= 25 − 1) belong to the block statistics assumed to exist. The reduction in this

sum of squares caused by fitting the statistics b_n, b_p, and b_k is found, as usual, by means of the relation

$$Sy'^2 = b_nSn^2 + b_pSp^2 + b_kSk^2 = 209.33 + 19.95 + 199.07 = 428.35, \quad (29)$$

corresponding to 3 degrees of freedom, since it corresponds to the three statistics so fitted. The reduced sum of squared deviations is, therefore, $Sd^2 = S(Y - Y')^2 = 740 - 428.35 = 311.65$, corresponding to 22df.

The estimate of the variance, V, is, therefore, $V = 311.65/22 = 14.1659$. The variance of b_n is then $C_{aa}V = (0.0935088)14.1659 = 1.3246$. The standard error of b_n = S.E.$_{b_n}$ = $(1.3246)^{1/2} = 1.151$, and the corresponding value of $t = 3.1479/1.151 = 2.73$, which is significant at the 5% point; since the value of t at the 5% for 22df. is 2.074.

TABLE IV

Analysis of the total sum of squared deviations

Source of the deviations	Degrees of freedom	Sum of squared deviations	Variance estimate	F values		
				Experimental	5%	1%
Total	49	4,651				
Blocks	4	802				
Treatments	9	922	102.4	1.26	2.15	2.94
Error	36	2,927	81.31			

Similarly, V_{b_p} = 1.9784, S.E.$_{b_p}$ = 1.407, and t_{b_p} = 0.37, which is not significant. For b_k, the corresponding figures are: V_{b_k} = 0.5609, S.E.$_{b_k}$ = 0.7489, and t_{b_k} = 2.67, which is also significant at the 5% point.

The performed analysis indicates that the crop responded significantly to the applications of nitrogen and potash, whereas it did not do so with respect to those of phosphorus.

The usual analysis of variance of the yield data of the experiment under study yielded table IV, the first block consisting of plots 1 to 10, the second block of plots 11 to 20, etc.

The usual hypothesis thus fails to show any significant effect of the treatments on the yields. However, since in the previous analysis but 3 statistics were fitted to explain the effects of the fertilizer applications whereas in this analysis 9 such statistics were fitted, a new analysis was made in an attempt to explain the sum of squares due to treatments in this last analysis by the use of only 3 statistics. Such an attempt indicated that of the sum of squares due to treatments, 922, a total of 334 could be explained by fitting the 3 statistics, there remaining 588 to be assigned to interactions of one sort or another. If this remainder is pooled with the sum of squares due to error, 2,927, a total of 3,515 is obtained, subject to

36 + 6 = 42df. The new estimate of the error variance would be, therefore, 83,69, of the same order of magnitude as the value formerly obtained of 81.31. On testing the fitted statistics, whose values came out to be: $b_n = 2.167$, $b_p = 0.529$ and $b_k = 1.188$, it is found that none of them is significant.

On comparing the variance estimate obtained by the last method of analysis, 83.69, with that obtained by the proposed method, 14.1659, one obtains a relative precision of 5.9078, (= 83.69/14.1659), in favor of the 2-plot blocks hypothesis as against the usual hypothesis, in this case the 10-plot blocks hypothesis. There has been, in this case, therefore, an increase in precision of 491 per cent due to the change in hypothesis.

The results of this experiment might have been also interpreted by assuming 5-plot incomplete blocks consisting of either the plots receiving treatments *A*, *C*, *E*, *G* and *I* or treatments *B*, *D*, *F*, *H*, and *J*. The results of such an analysis appear in table V.

TABLE V

Analysis of the total sum of squared deviations

Source of the deviations	Degrees of freedom	Sum of squared deviations	Variance estimate	*F* values		
				Experimental	5%	1%
Total	49	4,651				
Blocks	9	2,317				
Treatments	8	830	103.75	2.21	2.25	3.12
Error	32	1,504	47.00			

The observed differences between the treatment means are again not significant, although there has been an increase in precision of 81.31/47 − 100% = 73%, by the use of the smaller 5-plot blocks.

In case that an experiment consists of an odd number of plots, either by design or accident, the contribution of all but 3 plots to the numerical coefficients of the squares and products which appear in Sd^2 will be calculated as described above in the case of table II, by finding the respective differences between the adjacent plots, finding all possible squares and products of these differences, adding these squares and products, and dividing each of these sums by 2.

The contribution of the 3 plots left out in the previous calculations may be computed as follows: The set of differences corresponding to the first of these plots will be formed by subtracting the sum of the corresponding figures of the second and third plot from twice the corresponding figures of the first plot. The set of differences corresponding to the second plot will be likewise formed by subtracting the sum of the corresponding figures of the first and third plot from twice the corresponding figures of the

second plot. In a similar way are found the differences corresponding to the third plot.

Again the sums of all possible squares and products of these sets of differences are found, but in this case the sums must be divided by 9 in order to determine the contributions of these last 3 plots to the numerical coefficients. The sum of the contributions of both sets of plots will give the required numerical coefficients of the squares and products which appear in Sd^2.

To illustrate the way of finding the contribution of these last 3 plots in such an analysis, one may assume that plots 48, 49, and 50 of diagram I are the remaining 3 plots in the case of an experiment with an uneven number of plots. The value of d_Y for plot 48 would be found as follows: Since plot 48 received treatment E, with 2 units of nitrogen; plot 49 received treatment C, with no nitrogen; and plot 50 received treatment A, with no nitrogen, the respective difference, d_N, would be, therefore, $2(2) - 0 - 0 = 4$. The corresponding d_P would be $2(1) - 1 - 0 = 1$; the corresponding d_K would be $2(3) - 4 - 0 = 2$; and the corresponding d_Y would be $2(59) - 37 - 30 = 118 - 67 = 51$. In a similar way, the differences corresponding to plots 49 and 50 are found.

In the above discussion, the question of the significance of the difference between any pair of statistics fitted to data of the kind mentioned has been dispensed with. In the cases where several varieties or field practices have been tested, the interest centers, not on whether a given variety or practice produces a significant departure of the measured effect from the mean effect of all the varieties or practices tested, but on whether the given variety or practice causes a significantly greater or smaller effect than some other variety or practice included in the test. In such cases, the number of statistics fitted to the data is one less than the number of treatments under study. Thus, if, say, varieties A, B, C, D, and E were tested in one of such experiments, statistics for A, B, C, and D only would be fitted, it being understood that the statistic for E is equal to the negative sum of the statistics corresponding to the other four varieties, i.e., $E = -(A + B + C + D)$. The corresponding C_{ei}'s would be found by means of the relations:

$$C_{ea} = -(C_{aa} + C_{ab} + C_{ac} + C_{ad}), \tag{30}$$

$$C_{eb} = -(C_{ab} + C_{bb} + C_{bc} + C_{bd}), \tag{31}$$

$$C_{ec} = -(C_{ac} + C_{bc} + C_{cc} + C_{cd}), \tag{32}$$

$$C_{ed} = -(C_{ad} + C_{bd} + C_{cd} + C_{dd}), \tag{33}$$

and

$$C_{ee} = -(C_{ae} + C_{be} + C_{ce} + C_{de}). \tag{34}$$

The variance estimate of a difference between any two such statistics, say A and B, would be found by means of the relation $(C_{aa} - 2C_{ab} + C_{bb})V$,

from which the corresponding values of the standard error and t would be found accordingly.

It now remains to present a resumé of the results obtained to date in the use of the 2-plot blocks method in the interpretation of the results of field trials. It may be stated that, in general, and in full accord with the theoretical foundations of the method, it is most useful as compared with

TABLE VI

Comparison of efficiency of the 2-plot blocks method with the usual analysis of "variance" method of interpreting the results of complete randomized blocks experiments

Crop	Nature of test	Factor studied	Coefficient of variability obtained by usual analysis of "variance" method	Increase in efficiency due to the use of the 2-plot blocks method
			%	%
Corn, 1st crop	Fertilizer	Yield	11.04	13.15
Corn, 2nd crop	"	"	17.13	−6.43
Corn, 3rd crop	"	"	9.97	−11.00
Alfalfa, 16 cuttings	" unlimed	"	21.93	490.81
" , " "	" limed	"	7.13	32.60
Beans	"	"	16.79	27.86
Sweet potato, 1st crop	"	"	43.71	248.75
" " , 2nd "	"	"	26.92	51.29
Cucumbers, 1932–33	"	"	18.38	11.31
Cucumbers, 1933–34	"	"	18.20	109.32
" , 1935–36	"	"	27.67	74.88
" , 1936–37	"	"	21.02	75.15
" , 1941–42	"	"	25.25	24.81
Cotton	"	"	16.27	−7.61
Sugar cane, Aguirre 1942–43	Varietal	Yield of cane	17.27	53.48
" " "	"	Tons sugar/A	17.31	56.93
" " "	"	Sugar % cane	4.62	−13.18
Sugar cane, Río Piedras 1941–42	"	Tons sugar/A	27.60	190.16

the usual method of analysis in cases where soil heterogeneity affects markedly the effect under study, the more so the more variable the soil. In table VI, the results of comparisons of the 2-plot blocks method with the usual method of analysis are presented. In every case the comparisons between the two methods have been made by using the same number of statistics to explain the effects of the treatments.

As seen above, in by far the majority of the cases, the 2-plot blocks method has been more precise than the complete randomized blocks

method. The method, however, takes more time and work to apply to the results of trials not specially designed for its application. By using some sort of balanced design, however, the work of calculation of the results of experiments by means of the proposed method may be considerably shortened.

Thus, in testing 6 treatments, each of the treatments might be included with each of the other ones an equal number of times in the 2-plot blocks. This would require replicating each treatment either 5 times or some multiple of 5. In general, this would require $m(n-1)$ replications where n is the number of treatments tested and m is any positive integer. Such a layout for 6 treatments replicated 5 times could be something as illustrated in diagram II.

DIAGRAM II

Layout for a plot distribution to test 6 treatments (A, B, C, D, E, and F) each replicated 5 times

AB	CD	EF
EC	FA	BD
FD	BC	AE
CA	DE	FB
BE	CF	DA

The advantage of such a balanced design would be that all C_{ii}'s would be equal, i.e., $C_{aa} = C_{bb} = C_{cc} = C_{dd} = C_{ee} = C_{ff}$, and, similarly, all C_{ij}'s would be equal, i.e., $C_{ab} = C_{ac} = \cdots = C_{ef}$. Therefore, the same variance estimate would be used to test each effect constant and, likewise, there would be but one variance estimate to test the difference between any two such effect constants. More information relative to the possibilities of the use of these balanced designs will be given in another article to be published shortly after this one in this same Journal. In said article, the manuscript of which has been already prepared, the modified calculational technique is discussed and a numerical example is presented.

REFERENCES

1. Fisher, R. A., Statistical Methods for Research Workers, 4th edition, revised and enlarged. Oliver and Boyd, Edinburgh. 1932.
2. Mills, F. C., Statistical Methods, Revised. Henry Holt and Co., New York. 1939.
3. Rider, P. R., Modern Statistical Methods. John Wiley and Sons, Inc., New York. 1939.
4. Snedecor, G. W., Statistical Methods, 3rd edition. The Iowa State College Press, Ames, Iowa. 1940.

A NEW METHOD OF PERFORMING FIELD TRIALS

By BERNARDO G. CAPÓ

Biometrician

INTRODUCTION

In a recent article (1) in which the author discussed a new method of interpreting the results of field trials, he presented evidence tending to show that the assumption of a different effect constant for every pair of adjacent plots of a randomized block field trial led to a greater precision in the reduction of the corresponding data than the assumption of a different effect constant for every different complete block of the experiment. In said article he presented a method which might be used in this connection in order to reduce the calculational work to a minimum for the interpretation of such experiments by use of the proposed new assumption. The work involved in the application of said method to the results of randomized block experiments already performed, however, is greater than the work required for the application of the usual methods of statistical analysis in regard to which said experiments have been designed. On the other hand, the increase in precision obtained by the application of the suggested method seems to the author to warrant fully the increase in such calculational work.

As it was also mentioned in the article referred to, it is possible to reduce the number of operations required for the application of that method to the interpretation of field trials by designing such field trials so as to make use of the proposed method. The purpose of this article is to suggest designs for possible use in this connection as well as to discuss the procedure to be followed in the interpretation of such specially designed experiments.

REQUIREMENTS TO BE FULFILLED BY FIELD TRIALS

Field trials must fulfill certain requirements in order that the application of the proposed method to the interpretation of their results be most economical in time and effort. These requirements are as follows: 1—As many different groups of two treatments each must be made as is possible. 2—The experimental field must be divided into blocks of 2-adjacent plots each, so that the topographies of the two plots in any given block will be as nearly alike as possible. 3—Each group of two treatments must be assigned at random to be tested in one of the 2-plot blocks. The plot of the respective 2-plot block in which one or the other of the two treatments of the group assigned to it will be tested must be also decided at random. 4—The same number of replications of each group of treatments must be used. That is,

if n treatments are to be tried, there should be $m(n - 1)$ replications, where m is any positive integer. Thus, if 6 treatments are to be compared, either 5 or 10 or 15, etc. replications of the treatments are required.

EXAMPLES OF DESIGNS OF POSSIBLE USE

In order to give illustrations of the possibilities of this type of layout, two examples will be given of possible different designs for a trial in which 5 treatments are repeated 4 times each.

Where 5 treatments, say A, B, C, D, and E, are to be repeated 4 times each, the experiment should include 10 groupings as follows: AB, AC, AD, AE, BC, BD, BE, CD, CE, and DE. The 20-plot experimental field is divided into ten 2-plot blocks, and the above treatment pairs are

DIAGRAM I

Possible layouts of plots

(a)		(b)		
AB	BC	AB	EC	D
EC	EA	ED	CB	A
AD	BE	AE	BD	C
ED	AC	CA	EB	D
DB	CD			

assigned at random to be tested in one or another of the 10 blocks, there being also randomization of the order of the treatments to be tested within each block. A possible geographical distribution might be as indicated by Diagram I (a).

Another possible arrangement, which lends itself to possible interpretation by either the 2-plot block method or by the ordinary method of analysis for a 4-randomized-block trial, is shown in Diagram I (b).

THEORY OF THE PROCEDURE OF CALCULATION

As indicated in the article (1) previously referred to, the regression equation to be used in explaining the results of such an experiment would be

$$Y_i = B_i + T_i, \qquad (1)$$

where Y_i is the observed effect, say yield, of plot i, B_i is the effect constant of the block of which plot i forms part, and T_i is the effect constant of the treatment tested in plot i. After writing down the respective equation

for each of the plots used, the problem becomes one of finding the most probable values of the B_i's and the T_i's, and the standard errors of the latter for comparison with one another. It was suggested in said article that a convenient way to do this would be one in which the differences between the corresponding figures of the paired plots were found and then, following the usual methods of regression analysis, determining the most probable values of the T_i's and their standard errors.

Now, if the experiment has been carried out in such a way that every treatment has been thus paired with every other treatment *once*, it will be found that by the application of the above procedure, the value of each of the C_{ii}'s will come out to be $2(n-1)/n^2$ and the value of each of the C_{ij}'s $(i \neq j)$ will come out to be $-2/n^2$ where n is the number of treatments tested.

By knowing then *Say*, *Sby*, *Scy*, *Sdy*, *Sey*, and *Syy*, and by the proper use of the C_{ii}'s and C_{ij}'s, already known, the values of the treatment effect constants, i.e., A, B, C, D, and E, may be found together with their standard errors. One may thus test the significance between the estimated difference in effects between any two of the treatments tested.

The theory of the procedure outlined in the last paragraph is so well known to those interested in it, that it does not merit further consideration. It may be worthwhile, however, to derive the relations $C_{ii} = 2(n-1)/n^2$ and $C_{ij}(i \neq j) = -2/n^2$. This will be done now.

Let it be assumed that n treatments have been tested in a trial of the kind under consideration. Let the effects of the different treatments above or below the mean effect of all treatments be designated by $A, B, C, \cdots, N$, where N is the effect of nth treatment. The effect equations of the type

$$Y_i = B_i + T_i$$

may then be written for each plot.

Since any given treatment, say A, has been replicated $n-1$ times, there will be $n-1$ such equations in which the treatment effect A appears. Also, on account of this, there will be $n-1$ equations formed by finding the differences between the equations corresponding to the treatment A plots and those corresponding to the other plots paired to said treatment A plots. These equations might be written as:

$$\begin{aligned}
\pm A \mp B &= d_{\pm A \mp B} \\
\pm A \mp C &= d_{\pm A \mp C} \\
\pm A \mp D &= d_{\pm A \mp D} \\
&\cdots \\
\pm A \mp (N-1) &= d_{\pm A \mp (N-1)} \\
\pm A \mp N &= d_{\pm A \mp N}
\end{aligned}$$

where the d's stand for the respective differences between the effects noticed in the A plots as against the effects noticed in the plots paired with the A plots. Since, however, $A, B, C, \cdots$, and N have been assumed to be the effects of the treatments above or below the mean effect of all the treatments,

$$A + B + C + \cdots + N = 0,$$

or

$$N = -(A + B + C + \cdots + (N - 1)).$$

The last of the above equations, therefore, should appear rather as $\pm 2A \pm B \pm C \pm \cdots \pm (N - 1) = d_{\pm A \mp N}$, due to the relation for N previously pointed out. Further, the effect constant A would also appear in all the equations corresponding to comparisons between plots receiving treatment N and anyone of the other treatments.

In other words, A appears with a coefficient of ± 1 in the $n - 2$ equations corresponding to treatment A comparisons with treatments $B, C, D, \cdots, N - 1$, and also in the $n - 2$ equations corresponding to the comparisons of treatments $B, C, D, \cdots$, and $N - 1$ with treatment N; and furthermore, with a coefficient of ± 2, in the equation corresponding to its comparison with treatment N. That is, A will appear with a coefficient of ± 1 in $2n - 4$ equations and with a coefficient of ± 2 in one equation. This is true also of each of the other treatment effect constants.

Therefore, the sum of the *squares* of the A coefficients will be $(2n - 4) + 4 = 2n$. This is true also of the sum of the squares of the coefficients of any other treatment effect constant. But since, as was demonstrated in the former article, on finding the sums of squares and products of the differences between the coefficients of the treatment effect constants in the original equations, one would be finding really $2Sa^2$, $2Sab, \cdots, 2Sa(n - 1)$ instead of Sa^2, $Sab, \cdots, Sa(n - 1)$ which are the figures sought, the above value of $2n$ must be divided by 2, so that the value of $Sa^2 = Sb^2 = \cdots = S(n - 1)^2 = n$.

Now, in the equation corresponding to the comparison of A with B, one of the treatment effect constants is $+1$ and the other -1, so that their product is -1. In the equation corresponding to the comparison of A with N, the coefficient of A will be ± 2 and that of B will be ± 1, so that the products of these coefficients will be $+2$. This is true also of the equation corresponding to the comparison of B with N, where the coefficient of B is ± 2 and the coefficient of A is ± 1. In each of the $n - 3$ equations corresponding to the comparisons of each of the other treatments with N, the coefficients of both A and B will be ± 1, so that their product in each case will be ± 1. Therefore, the sum of the products of the coefficients of

A and B in these equations will be $-1 + 2 + 2 + (n - 3) = n$. Again, this must be divided by 2 in order to get the proper value of Sab. Thus, $Sab = Sac = \cdots = Sa(n-1) = Sbc = \cdots = S(n-2)(n-1) = n/2$.

The normal equations will be, therefore,

$$\begin{aligned}
An + Bn/2 + Cn/2 + \cdots + (N-1)n/2 &= Say, \\
An/2 + Bn + Cn/2 + \cdots + (N-1)n/2 &= Sby, \\
\cdot \quad\quad \cdot \quad\quad \cdot \quad\quad\quad\quad \cdot \quad\quad \cdot & \\
\cdot \quad\quad \cdot \quad\quad \cdot \quad\quad\quad\quad \cdot \quad\quad \cdot & \\
\cdot \quad\quad \cdot \quad\quad \cdot \quad\quad\quad\quad \cdot \quad\quad \cdot & \\
An/2 + Bn/2 + Cn/2 + \cdots + (N-1)n &= S(n-1)y,
\end{aligned}$$

giving rise to the $(n-1)$ columned determinant

$$D_n = \underbrace{\begin{vmatrix} n & n/2 & n/2 & \cdots & n/2 \\ n/2 & n & n/2 & \cdots & n/2 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ n/2 & n/2 & n/2 & \cdots & n \end{vmatrix}}_{(n-1) \text{ columns}}$$

from which the values of C_{aa}, C_{ab}, $\cdots$, $C_{(n-1)(n-1)}$ may be evaluated as follows.

By subtracting the last row in D_n from each of the other rows, adding the first to the last column in the expression thus obtained, expanding the expression thus obtained by minors along the first row, and by continuing such additions of the first to the last column and expansions by minors along the first row in the resulting expressions, one obtains finally

$$D_n = \left(\frac{n}{2}\right)^{n-3} \begin{vmatrix} \frac{n}{2} & 0 \\ \frac{n}{2} & n + (n-2)\frac{n}{2} \end{vmatrix} = \left(\frac{n}{2}\right)^{n-2} \left[n + (n-2)\frac{n}{2}\right] = n\left(\frac{n}{2}\right)^{n-1}.$$

Now, $C_{aa} = N_{aa}/D_n$, where

$$N_{aa} = \underbrace{\begin{vmatrix} n & n/2 & \cdots & n/2 \\ n/2 & n & \cdots & n/2 \\ \cdot & & & \\ \cdot & & & \\ n/2 & n/2 & \cdots & n \end{vmatrix}}_{(n-2) \text{ columns}}$$

The evaluation of N_{aa}, following a procedure identical to that used in evaluating D_n, yields the expression

$$N_{aa} = \left(\frac{n}{2}\right)^{n-4} \begin{vmatrix} n/2 & 0 \\ n/2 & n + (n-3)\frac{n}{2} \end{vmatrix}$$

$$= \left(\frac{n}{2}\right)^{n-3} \left[n + (n-3)\frac{n}{2}\right] = (n-1)\left(\frac{n}{2}\right)^{n-2}.$$

C_{aa} is, therefore, equal to $2(n-1)/n^2$. On finding the corresponding expressions for the other treatments, it is found that $N_{aa} = N_{bb} = \cdots = N_{(n-1)(n-1)}$ and, therefore, $C_{aa} = C_{bb} = \cdots = C_{(n-1)(n-1)} = 2(n-1)/n^2$.

Now, $C_{ab} = N_{ab}/D_n$, where

$$-N_{ab} = \underbrace{\begin{vmatrix} n/2 & n/2 & \cdots & n/2 \\ n/2 & n & \cdots & n/2 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ n/2 & n/2 & & n \end{vmatrix}}_{(n-2) \text{ columns}}$$

This determinant may be evaluated by subtracting the first row from each of the other ones and expanding the resulting expression by minors along the first column and continuing with expansions of the same kind successively with the resulting expressions, obtaining finally $\left(\frac{n}{2}\right)^{n-2}$. Thus, $N_{ab} = -\left(\frac{n}{2}\right)^{n-2}$, and C_{ab} comes out to be $-2/n^2$. The determinants from which $N_{ac}, N_{ad}, \cdots$, and $N_{(n-2)(n-1)}$ are to be evaluated will differ in every case from that used to evaluate N_{ab}, but by proper inversions of rows and columns they can be brought to coincide with the above one, so that $N_{ac} = N_{ad} = \cdots = N_{(n-2)(n-1)}$ and therefore, $C_{ab} = C_{ac} = \cdots = C_{(n-2)(n-1)} = -2/n^2$.

In experiments of this nature where each treatment is repeated $2(n-1)$ times, N_{ii} is $(n-1)\left(\frac{n}{2}\right)^{n-2}2^{n-2} = (n-1)(n)^{n-2}$, $N_{ij} = -\left(\frac{n}{2}\right)^{n-2}2^{n-2} = -(n)^{n-2}$, and $D_n = n\left(\frac{n}{2}\right)^{n-1}(2)^{n-1} = n(n)^{n-1}$. Therefore, $C_{ii} = \frac{(n-1)n^{n-2}}{n(n)^{n-1}} = \frac{n-1}{n^2}$, and $C_{ij} = \frac{-(n)^{n-2}}{n(n)^{n-1}} = \frac{-1}{n^2}$. That is, as the sums of squares and products double, the values of the C_{ii}'s and C_{ij}'s are divided

by 2. If the treatments are repeated $3(n-1)$ times, $C_{ii} = \frac{1}{3}\left(\frac{2(n-1)}{n^2}\right)$ and $C_{ij} = \frac{1}{3}\left(\frac{-2}{n^2}\right)$, and similarly for other multiples of $n-1$ replications.

In the above discussion no reference has been made to C_{nn} and to the C_{in}'s, that is, the values corresponding to the last variety. Now $C_{nn} = C_{aa} = C_{bb} = \cdots = C_{(n-1)(n-1)}$ and $C_{in} = C_{ab} = C_{ac} = \cdots = C_{(n-2)(n-1)}$. This may be proved as follows: $N = -(A + B + C + \cdots + (N-1))$, and therefore, for $n-1$ replications, $C_{an} = -(C_{aa} + C_{ab} + C_{ac} + \cdots + C_{a(n-1)}) = -C_{aa} - (n-2)C_{ab} = \frac{-2(n-1)}{n^2} - (n-2)\left(\frac{-2}{n^2}\right) = \frac{-2}{n^2} = C_{ab} = C_{ac} = \cdots = C_{(n-2)(n-1)}$. This holds also for $C_{bn}, C_{cn}, \cdots, C_{(n-1)n}$. Also, $C_{nn} = -(C_{an} + C_{bn} + \cdots + C_{(n-1)n}) = -(n-1)\left(\frac{-2}{n^2}\right) = 2(n-1)/n^2 = C_{aa}$, etc. The above relations hold also for other multiples of $(n-1)$ replications.

Numerical example

Diagram II shows the geographical distribution of the plots, together with the treatments and yields per plot, in cwts. per acre, of a fertilizer experiment with cotton performed by Messrs. A. Riollano and J. Pastor Rodríguez in cooperation with the author, on Coto clay at Mantilla Farm, Isabela.

Column (1) of table I shows one way in which the plots may be paired so that one may employ the procedure of calculation discussed above. As will be noticed, the paired plots lie side by side in all but two cases: where plot 56 has been paired to plot 46, and where plot 35 has been paired to plot 47. In these two cases the topography of the field in the section where these plots were located suggested that they should be paired. Columns (2) to (8) show the coefficients of the treatment constants corresponding to each of the paired plots. The constant of treatment H has been expressed in terms of the constants of the other seven treatments according to the relation $H = -(A + B + C + D + E + F + G)$. Column (9) shows the respective differences between the yields of the paired plots, taken in the order indicated in column (1). Though unnecessary, the plots of each pair were paired so that these differences in yields were positive. The numbers in the central portions of columns (10) to (16) are the products of the entries in columns (2) to (8) by those of column (9), as expressed at the tops of the respective columns. The first of the 2 rows at the bottom of columns (10) to (16) consists of the sums of the

entries in the body of each of the respective columns. The entries of the second row at the bottom of these columns are each one-half of the entries of the first row and indicate the values of *Say*, *Sby*, $\cdots$, *Sgy*, according

DIAGRAM II

Geographical distribution of plots of fertilizer experiment with cotton

52 D 5.0	53 A 6.1	54 B 6.3	55 F 7.6	56 C 9.5		
48 H 4.2	49 C 5.0	50 H 6.6	51 A 5.5			
42 F 8.6	43 G 6.7	44 E 5.9	45 G 5.2	46 B 8.0	47 D 7.8	
36 E 7.3	37 H 3.4	38 D 5.5	39 F 7.1	40 E 8.3	41 H 3.5	
29 C 6.4	30 G 6.4	31 B 7.3	32 C 6.2	33 A 4.0	34 G 8.5	35 B 8.4
22 D 7.2	23 A 5.3	24 F 6.6	25 G 8.3	26 D 6.5	27 E 8.5	28 C 8.7
15 E 6.8	16 F 7.2	17 H 3.8	18 B 8.1	19 H 5.5	20 A 5.6	21 F 8.4
8 C 4.6	9 D 7.6	10 A 4.0	11 G 7.7	12 F 8.7	13 B 9.9	14 E 7.0
1 B 6.5	2 E 5.2	3 G 7.0	4 D 8.1	5 A 5.5	6 C 7.2	7 H 6.0

to whether the heading of their respective columns are *AY*, *BY*, $\cdots$, *GY*. The theoretical reason for this procedure of finding *Say*, *Sby*, $\cdots$, *Sgy* has been given elsewhere (1).

Now, in this experiment each of the 8 treatments has been paired once

TABLE I

Calculation of Say, Sby, ···, Sgy

(1)	(2)	(3)	(4)	(5)	(8)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Plots	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*Y*	*AY*	*BY*	*CY*	*DY*	*EY*	*FY*	*GY*	*YY*
1–2		1			−1			1.3		1.3			−1.3			1.69
4–3				1			−1	1.1				1.1			−1.1	1.21
6–5	−1		1					1.7	−1.7		1.7					2.89
14–7	1	1	1	1	2	1	1	1.0	1.0	1.0	1.0	1.0	2.0	1.0	1.0	1.00
15–8			−1		1			2.2			−2.2		2.2			4.84
9–16				1		−1		0.4				0.4		−0.4		0.16
10–17	2	1	1	1	1	1	1	0.2	0.4	0.2	0.2	0.2	0.2	0.2	0.2	0.04
18–11		1					−1	0.4		0.4					−0.4	0.16
12–19	1	1	1	1	1	2	1	3.2	3.2	3.2	3.2	3.2	3.2	6.4	3.2	10.24
13–20	−1	1						4.3	−4.3	4.3						18.49
28–21			1			−1		0.3			0.3			−0.3		0.09
27–26				−1	1			2.0				−2.0	2.0			4.00
22–29			−1	1				0.8			−0.8	0.8				0.64
30–23	−1						1	1.1	−1.1						1.1	1.21
31–24		1				−1		0.7		0.7				−0.7		0.49
25–32			−1				1	2.1			−2.1				2.1	4.41
40–33	−1				1			4.3	−4.3				4.3			18.49
34–41	1	1	1	1	1	1	2	5.0	5.0	5.0	5.0	5.0	5.0	5.0	10.0	25.00
35–47		1		−1				0.6		0.6		−0.6				0.36
42–36					−1	1		1.3					−1.3	1.3		1.69
38–37	1	1	1	2	1	1	1	2.1	2.1	2.1	2.1	4.2	2.1	2.1	2.1	4.41
43–44					−1		1	0.8					−0.8		0.8	0.64
39–45						1	−1	1.9						1.9	−1.9	3.61
56–46		−1	1					1.5		−1.5	1.5					2.25
49–48	1	1	2	1	1	1	1	0.8	0.8	0.8	1.6	0.8	0.8	0.8	0.8	0.64
53–52	1			−1				1.1	1.1			−1.1				1.21
50–54	−1	−2	−1	−1	−1	−1	−1	0.3	−0.3	−0.6	−0.3	−0.3	−0.3	−0.3	−0.3	0.09
55–51	−1					1		2.1	−2.1					2.1		4.41
									−0.2	17.5	11.2	12.7	18.1	19.1	17.6	114.36

with each of the other treatments, and as mentioned in the section on the theory of the procedure of calculation,

$$C_{aa} = C_{bb} = \cdots = C_{gg} = C_{hh} = +2\,(8-1)/8^2 = 14/64 = 7/32,$$

and

$$C_{ab} = C_{ac} = \cdots = C_{bc} = C_{bd} = \cdots = C_{gh} = -2/8^2 = -2/64 = -1/32.$$

Knowing, then, the C_{ii}'s and the C_{ij}'s, together with Say, Sby, $\cdots$, Sgy, one can calculate the yield constants corresponding to each treatment and test the significance of differences between these constants.
Thus,

$$\begin{aligned}
A = b_A &= C_{aa}Say + C_{ab}Sby + C_{ac}Scy + C_{ad}Sdy \\
&\quad + C_{ae}Sey + C_{af}Sfy + C_{ag}Sgy \\
&= (7Say - Sby - Scy - Sdy - Sey - Sfy - Sgy)/32, \\
&= (8Say - (Say + Sby + Scy + Sdy + Sey + Sfy + Sgy))/32 \\
&= (-0.8 - 48.00)/32 = -48.8/32 = -1.525. \\
B = b_B &= (8Sby - 48.00)/32 = (70.00 - 48.00)/32 = 22/32 = 0.6875 \\
C = b_C &= (8Scy - 48.00)/32 = (44.80 - 48.00)/32 \\
&= -3.2/32 = -0.1000 \\
D = b_D &= (8Sdy - 48.00)/32 = (50.80 - 48.00)32 = 2.8/32 = 0.0875 \\
E = b_E &= (8Sey - 48.00)/32 = (72.40 - 48.00)/32 = 24.4/32 = 0.7625 \\
F = b_F &= (8Sfy - 48.00)/32 = (76.40 - 48.00)/32 = 28.4/32 = 0.8875 \\
G = b_G &= (8Sgy - 48.00)/32 = (70.40 - 48.00)/32 = 22.4/32 = 0.7000 \\
H &= -(A + B + C + D + E + F + G) = -1.5000.
\end{aligned}$$

To make comparisons between these statistics, one must calculate their standard errors. The sum of the squared deviations of the yields corrected for differences in fertility between the different 2-plot blocks is found by dividing the sum of the squares of the entries in column (9) of table I by 2. The squares of the corresponding entries constitute the body of column (17) of table I, their sum is the next to the last item in that column, and the required sum of squared deviations corrected for differences in fertility between the 2-plot blocks is one-half the next to the last item in column (17), that is, 57.18. This sum of squared deviations is subject to 28 degrees of freedom, since 28 degrees of freedom have been lost: one *df* corresponding to the mean of all the yields and 27 *df* due to the fitting of the 27 different block constants necessary to express the relations between the fertilities of the 28 2-plot blocks.

The reduction of the sum of squared deviations due to the fitting of the treatment constants to the yield data is, according to the usual relation, $Sy'y' = ASay + BSby + \cdots + GSgy = (-1.525)(-0.1) + 0.6875\ (8.75) + \cdots + 0.7000\ (8.80) = 27.70$, subject to 7 *df*.

The reduced sum of squared deviations is, therefore, $S(Y - Y')^2 = Syy - Sy'y' = 57.18 - 27.70 = 29.48$, subject to $28 - 7 = 21$ *df*. The estimate of the reduced variance, V', is then $29.48/21 = 1.4037$.

The reduction in the sum of squared deviations due to the use of the 7 treatment constants may be tested by means of the F-test as follows: $F = 27.70/7(1.4037) = 27.70/9.8259 = 2.82$. Since the value of F at the 5% point, for 7 and 21 *df* is 2.49, the reduction in the sum of squared deviations due to the fitting of the treatment constants is a significant one.

The total sum of squared deviations of the yields of this experiment is 134.29, and as found above, the sum of squared deviations of the yields corrected for differences in fertility between blocks is 57.18. The fitting of the block constants led therefore to a reduction of $134.29 - 57.18 = 77.11$ in the sum of squared deviations. These relations are indicated in table II.

TABLE II

Analysis of the total sum of squared deviations following the 2-plot block method

Source of the deviations	Degrees of freedom	Sum of squared deviations	Variance estimate	F values		
				Experimental	5%	1%
Total	55	134.29				
Blocks	27	77.11				
Treatments	7	29.48	4.21	2.82	2.49	3.65
Error	21	27.70	1.4037			

In testing by means of the t-test any difference between the fitted treatment constants, the variance of any such difference would be found by multiplying V' by the corresponding factor of the nature of $C_{ii} - 2C_{ij} + C_{jj}$. Since, in the case of this experiment all the C_{ii}'s are equal and all the C_{ij}'s are equal, the variance of the difference between any two treatment constants is $(7/32 - 2(-1/32) + 7/32)\ 1.4037 = 0.70185$. The standard error of any such difference is, therefore, $(0.70185)^{1/2} = 0.8378$, and any such difference, to be significant at the 5% point, must exceed $2.080(0.8378) = 1.74$ cwts. cotton per acre.

The treatments compared in this test are described in table III. The corrected mean yield of any given treatment is found by adding algebraically the respective treatment yield constant as found above to the mean yield of all plots, 6.64.

Since treatments E, F, and G differed only with respect to the rate of the K_2O applications, the comparisons of their corrected mean yields or, what is equivalent, of their treatment yields constants, will furnish information on the effect of the K_2O applications on the cotton yields. Such

comparisons indicate that the effect of the K_2O applications on the cotton yields was not significant since the largest difference between any two of said constants was only 0.1875, whereas such a difference would have to exceed 1.74 cwts./A. to be significant at the 5% point.

Treatments *C*, *D*, and *G* varied only with respect to the P_2O_5 applications. In this case also, none of the differences between the yield constants were significant although they were of a larger magnitude than in the case of the K_2O applications.

A similar comparison of the yield constants of treatments *A*, *B*, and *G* indicates that there was a significant increase in yield, caused by the application of at least 100 lbs. NH_3 per acre, since the difference between the yield constants of treatments *A* and *B* was 2.22 cwts. cotton per acre, and the corresponding difference between the yield constants of treatments

TABLE III

Description of treatments and corrected mean yields

Treatment letter	Lbs. of nutrient applied per acre			Treatment yield constant	Corrected mean yield (cwts./acre)
	NH_3	P_2O_5	K_2O		
A	0	200	200	−1.5250	5.11
B	100	200	200	0.6875	7.33
C	200	0	200	−0.1000	6.54
D	200	100	200	0.0875	6.73
E	200	200	0	0.7625	7.40
F	200	200	100	0.8875	7.53
G	200	200	200	0.7000	7.34
H	0	0	0	−1.5000	5.14

A and *G* was 2.23 cwts. cotton per acre, both differences being significant at the 5% point and nearly so at the 1% point. The difference between the yield constants of treatments *B* and *G* was not significant indicating that the second 100 lbs. NH_3 applied per acre did not affect significantly the yields already obtained with the first 100 lbs. NH_3 applied per acre.

The yield constant of treatment *H* was almost equal to that of treatment *A*, corroborating the conclusions drawn above that nitrogen was the only one of the three elements tested which affected the yields significantly.

This experiment might have also been interpreted by using only 3 constants to explain the variations in yield caused by the various treatments as described in the article (1) previously referred to. Such a study was made yielding the same results previously obtained, except that the partial regression coefficient of the yields on the nitrogen applications came out to be significant at the 1% point.

SUMMARY

A new method of performing field experiments with relatively small numbers of treatments is described. The requirement to be fulfilled by the layouts of such field tests is specified and examples of possible designs for a 5-treatment experiment are illustrated. The theory of the procedure of calculation is discussed and a numerical example of said calculations is furnished in connection with the interpretation of a fertilizer experiment performed with cotton.

REFERENCE

1. Capó, B. G. A method of interpreting the results of field trials. This Journal: **28** (1), 7–21. 1944.